

BRIEF REPORT

Falling Behind: Time and Expectations-Based Reference Dependence

Rahul Bhui
Harvard University

Although the structure of simple decision making is well described as the noisy accumulation of evidence over time, the influence of subjective preferences for time use in deliberation remains underexplored. Here, I show how reference dependence based on expected time use can affect deliberative time allocation. In a motion discrimination task, I provide some participants with an indication of how long the block will take to complete. Once they exceed this benchmark, they spend less time accumulating information, at the expense of forgone rewards. By incorporating reference dependence into the drift diffusion model, I quantify the impact of the reference point on preferences directly. Moreover, I uncover the time course of preferences by estimating the subjective value of time in a sliding window. Thus, I provide novel evidence on the impact of temporal reference points and illuminate the role of subjective preferences in deliberative behavior.

Keywords: drift diffusion model, decision making, time allocation, reference dependence, expectations

To make decisions, people take time to process information. Although deliberation over simple alternatives is well characterized as the sequential sampling of noisy information (Gold & Shadlen, 2007; Laming, 1968; Ratcliff, 1978; Stone, 1960), the impact of subjective preferences for time use in this process has remained underexplored. Preferences are typically challenging to quantify, but understanding their role

is important for explaining and predicting deliberative behavior. In this article, I study how deliberation is influenced by reference points of expected time use. A decision maker who spends longer than anticipated on a task may feel discontentment, reducing his motivation to continue working (Kőszegi & Rabin, 2006).

The purpose of deliberation is to reduce uncertainty about which option is better. Accordingly, decisions frequently involve an inverse relationship between speed and accuracy; we can make judgments that are fast but error-prone or slow but high quality. In simple problems, patterns of choices and response times as well as neural activity have been formally captured using sequential sampling models, which describe decision making as the noisy accumulation of evidence over time until a desired standard of confidence is met (see Forstmann, Ratcliff, & Wagenmakers, 2016; Ratcliff, Smith, Brown, & McKoon, 2016). Although these models generate a realistic speed–accuracy trade-off, their applications are often agnostic as to how agents negotiate this

This article was published Online First October 11, 2018.

Funding from the Social Sciences and Humanities Research Council of Canada and the Harvard Mind/Brain/Behavior Initiative is gratefully acknowledged.

Thanks to Alex Alekseev, Ben Bushong, Colin Camerer, Kyle Carlson, Matt Chao, John Clithero, Ben Gillen, Taisuke Imai, Ian Krajbich, Matt Shum, Gerelt Tserenjigmid, Ryan Webb, and seminar participants at Harvard, Caltech, Alberta, ESA 2015, and Yale Whitebox 2016 for helpful comments and discussions.

Correspondence concerning this article should be addressed to Rahul Bhui, Departments of Psychology and Economics & Center for Brain Science, Harvard University, 52 Oxford Street, Cambridge MA 02138. E-mail: rbhui@g.harvard.edu

tradeoff. From a theoretical standpoint, an agent's stopping criterion incorporates all costs and benefits of spending time (Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006), but empirical analyses of subjective preferences remain limited in scope. Subjective motivations are important drivers of deliberation but are also by nature difficult to observe. Mostly, studies have simply manipulated the speed–accuracy tradeoff using instructions that either ask participants to favor fast responding, or accurate responding, with straightforward qualitative results (Band, Ridderinkhof, & van der Molen, 2003; de Hollander et al., 2016; Forstmann et al., 2008; Hale, 1969; Herz et al., 2017; Howell & Kreidler, 1963; Ivanoff, Branning, & Marois, 2008; Osman et al., 2000; Palmer, Huk, & Shadlen, 2005; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998; Rinkenauer, Osman, Ulrich, Müller-Gethmann, & Mattes, 2004; van der Lubbe, Jaśkowski, Wauschkuhn, & Verleger, 2001; van Veen, Krug, & Carter, 2008; Zhang & Rowe, 2014). Can deliberative behavior be influenced in more sophisticated ways, and can we directly quantify the shifts in preferences that result?

Much research in behavioral economics demonstrates that outcomes are evaluated relative to reference points, which are increasingly thought to stem from a decision maker's expectations (Bell, 1985; Gul, 1991; Kőszegi & Rabin, 2006; Loomes & Sugden, 1986). Such theories help us intuitively and elegantly explain a range of empirical findings which are hard to understand using traditional assumptions (e.g., Bartling, Brandes, & Schunk, 2015; Eliaz & Spiegler, 2013; Pagel, 2017; Pope & Schweitzer, 2011). Here, I focus on reference dependence in the time dimension, which has been invoked to capture important economic regularities pertaining to labor supply (Kőszegi & Rabin, 2006; Crawford & Meng, 2011), but which has also been overshadowed by reference dependence in terms of money or objects. Intuitively, workers who exceed their expected shift length may be less willing to persevere. Time is a crucial part of deliberation, and so expectations of time use may naturally play a special role in this domain. Although subjective preferences are usually hard to measure, studying behavior in a setting where the structure of the decision process is well understood enables us to precisely quantify how preferences change.

I conduct a motion discrimination experiment in which some participants are provided with information about how long an entire block of trials should take to complete. Such participants speed up after exceeding this reference point, at the expense of reduced accuracy. Some evidence suggests that they also exhibit more displeasure as measured by their ratings of task satisfaction. To quantitatively assess changes in subjective preferences as directly as possible, I incorporate reference dependence into the most commonly used sequential sampling framework, the drift diffusion model. By estimating model parameters from trials within different temporal windows, I can infer the location and magnitude of the reference point's impact on preferences for time versus reward. This procedure reveals a steep jump in subjective costs at the reference point by a factor of nearly three. Hence, exceeding a temporal benchmark on block-level timescales may dramatically alter the willingness to continue deliberating. I thus provide novel evidence on the impact of time-based reference points and illuminate the role of subjective preferences in deliberative time allocation.

Theory

For well over a century, psychological research has made use of response times to help understand and predict behavior (Donders, 1869/1969; Hick, 1952; Jensen, 2006; Sternberg, 1966; Shepard & Metzler, 1971). Sequential sampling models were developed as computational tools to rigorously quantify insights about the cognitive processes that give rise to decisions (Busemeyer & Townsend, 1993; LaBerge, 1962; Pike, 1966; Shadlen & Newsome, 1996; Usher & McClelland, 2001; Vickers, 1970; Wang, 2002). These models propose that to arrive at a choice, noisy information about each alternative is continually sampled until confidence in one option or another reaches a threshold level (Gold & Shadlen, 2007). This framework provides a precise statistical account of how deeper cognitive parameters give rise to time spent and performance attained, and in so doing, allows for richer interpretations of behavioral data.

The earliest and most well-characterized sequential sampling model is the drift diffusion model (DDM), originally developed half a cen-

tury ago (Laming, 1968; Ratcliff, 1978; Stone, 1960). Beyond behavioral evidence showing that the DDM closely fits patterns of choice and response times in a variety of decision tasks, direct recordings of neural activity demonstrate that neurons in various brain regions implement evidence accumulation processes that match the model's structure (Gold & Shadlen, 2002; Hanes & Schall, 1996; Ratcliff, Cherian, & Segraves, 2003; Smith & Ratcliff, 2004; Shadlen & Newsome, 2001). Indeed, the basic functioning of neurons involves transmitting all-or-nothing signals that are triggered by inputs reaching a critical threshold. Although the DDM is commonly used to study perceptual choice (Gold & Shadlen, 2007; Piliastides, Ratcliff, & Sajda, 2006; Ratcliff et al., 2003; Ratcliff, Piliastides, & Sajda, 2009; Ratcliff & Rouder, 1998; Ratcliff & Smith, 2004; Smith & Ratcliff, 2004; Voss, Rothermund, & Voss, 2004), recent work extends it to value-based settings such as consumer purchasing decisions and intertemporal choice (Krajbich, Armel, & Rangel, 2010; Krajbich, Lu, Camerer, & Rangel, 2012; Krajbich & Rangel, 2011; Milosavljevic, Malmaud, Huth, & Rangel, 2010).

According to the DDM, the agent integrates evidence over time for one alternative or another until an evidence threshold is reached, and the corresponding decision is made. This accumulation includes inherent sensory noise and hence is modeled as a stochastic differential equation,

$$dx = A dt + c dW,$$

where $x(t)$ is the difference in evidence between the two alternatives (with $x(0) = 0$ in an unbiased decision), A is the accumulation or drift rate, and c represents the noise component. The change dx over the small time interval dt is broken up into the constant drift $A dt$ and the Gaussian white noise $c dW$ with mean 0 and variance $c^2 dt$. When the accumulated evidence x reaches the critical threshold $\pm z$, the corresponding choice is made.

This stochastic process generates a speed-accuracy trade-off regulated by the confidence threshold (z). A higher threshold entails a more stringent standard of evidence and reduced susceptibility to errors at the cost of greater decision time. Conversely, a lower threshold requires weaker evidence and thus less time to

make a decision but increases the error rate. The drift rate (A) and noise (c) parameters describe an individual's information processing faculties. Higher drift rates and lower accumulation noise mean superior performance in terms of higher accuracy rates with the same threshold. Key mathematical properties of individual performance conditional on the DDM parameters have been characterized. These properties come from solutions to the first passage problem in which the stochastic accumulation process crosses the decision threshold. In the current simple setup,¹ closed-form expressions exist for the accuracy, a , and mean decision time, t (Bogacz et al., 2006):

$$a(z, A, c) = \frac{e^{2Az/c^2}}{1 + e^{2Az/c^2}},$$

$$t(z, A, c) = \frac{z}{A} \tanh\left(\frac{Az}{c^2}\right).$$

The DDM was partly motivated by its origins in efficient statistical algorithms for hypothesis testing, appealing to researchers as an instance of ideal observer analysis, which has generally proven fruitful in psychophysics. Specifically, the DDM is the continuous sampling limit of the sequential probability ratio test, which minimizes response time for a given error rate in a Bayes optimal way (Arrow, Blackwell, & Girschick, 1949; Wald, 1947). Theoretical derivations of the DDM and some of its extensions imply that the decision threshold is selected to balance the benefit of spending more time working—which reflects increased chances of winning monetary payoffs—with the cost—which stems from the value of forgone leisure. Various experiments show that the estimated threshold and its neural representation are indeed modulated by many kinds of manipulations, including explicit instructions to emphasize either the speed or the accuracy of responses (Band et al., 2003; de Hollander et al., 2016; Forstmann et al., 2008; Hale, 1969; Herz et al., 2017; Howell & Kreidler, 1963; Ivanoff et al., 2008; Osman et al., 2000; Palmer et al., 2005; Ratcliff &

¹ Additional parameters dealing with, for example, variation in drift rate across trials are sometimes incorporated in the extended DDM.

Rouder, 1998; Ratcliff & McKoon, 2008; Rinkenauer et al., 2004; van der Lubbe et al., 2001; van Veen et al., 2008; Zhang & Rowe, 2014). These studies typically find that the instructions affect behavior via the decision threshold, though occasionally changes in other parameters are observed as well. However, these findings are rarely translated into their quantitative implications for preferences directly.

Formally, supposing each correct answer yields a payoff of wage w and time expenditure comes at an opportunity cost² of rate π , a decision maker is assumed to choose a threshold that maximizes expected utility $U(z; A, c) = wa(z; A, c) - \pi t(z; A, c)$. This optimization criterion is equivalent to the Bayes Risk criterion developed by Wald and Wolfowitz (1948) and Edwards (1965), which assumes decision makers minimize the cost function $BR = k_1 t + k_2(1 - a)$ for $k_1, k_2 > 0$ and is known to have a unique solution. The first-order condition with respect to z is

$$\frac{\partial U}{\partial z} = 0 = w \left(\frac{2Ae^{2Az^*/c^2}}{c^2(1 + e^{2Az^*/c^2})} - \left[\frac{z^*}{c^2} \operatorname{sech}\left(\frac{Az^*}{c^2}\right) + \frac{1}{A} \tanh\left(\frac{Az^*}{c^2}\right) \right] \right) \pi,$$

and rearranging this equation yields

$$\hat{\pi} = \frac{wA^2}{2Az^* + c^2 \sinh(2Az^*/c^2)}. \quad (*)$$

In this way the opportunity cost of time is identified for each person from the decision threshold (z), drift rate (A), and accumulation noise (c) parameters which can be estimated from individual accuracy and response time data.³

Theories of reference dependence imply a scaling up of costs when the reference point r is exceeded. A person with reference-dependent preferences will be displeased if he spends longer than expected on the task. That is, if he exceeds the reference point, a psychological tax applies to each additional moment of work. Expected utility is accordingly given by $U(z|r) = wa(z) - \lambda_{rp} \pi t(z)$ where λ_{rp} is equal to the degree of loss aversion $\lambda > 1$ after

passing the reference point and is 1 otherwise. The theory predicts that to mitigate the loss sensations, the decision maker will reduce time expenditure on the task. This reduction comes at the expense of accuracy, decreasing chances of monetary reward. In this setup, the inferred cost parameter in Expression (*) includes this extra λ coefficient after the reference point is passed (i.e., the right-hand side expression is an estimate of $\lambda\pi$ rather than just π). By comparing parameters in the early and late regimes, or across people given and not given reference points, the intensity of loss aversion can be isolated. I note that the exact psychological source of loss aversion could be multifaceted. Past work has used the term *reference dependence* in nonspecific ways that could include individual or social motivations. Though this is surely an important question in understanding the impact of reference dependence across domains, in the present work I remain agnostic about the precise nature of the aversion.

Method

The goal of the experiment was to test the empirical implications of temporal reference dependence, and quantitatively investigate its impact in terms of underlying subjective preferences. I used a standard perceptual paradigm, the random dot motion task, in which behavior is known to be closely fit by the drift diffusion model, allowing me to focus on economic preferences. In this task, many dots were moving on a screen; most were moving in random directions, but a proportion were moving consistently either left or right, and participants had to figure out that consistent direction. A tradeoff between speed and accuracy came into play. One could spend more time, gather more sensory information, and be more likely to answer correctly and get paid. Or one could spend less time on the task and have more leisure time afterward but face a higher chance of being wrong and forgoing payment. I manipulated the willing-

² Theoretically, the value of time is its opportunity cost, so I use these terms interchangeably.

³ It is worth observing that a more typical economic modeling approach would suppose that the agent directly chooses the amount of time, t , to spend on each trial of the task, yielding an accuracy of $a(t)$. However, this formulation would not allow estimation of preferences because the connection between observed accuracy and time depends endogenously on the individual's stopping criterion.

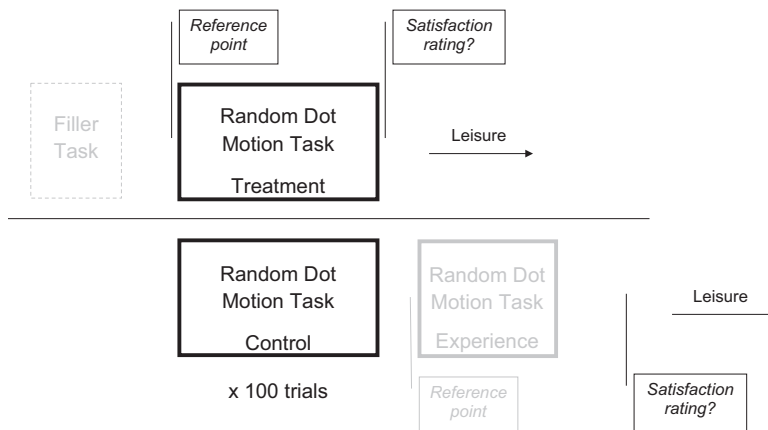


Figure 1. Setup of experiment. Participants engaged in blocks of perceptual decision-making tasks. In the treatment group, participants were provided with a temporal reference point indicating that a block of 100 dot motion trials should take about 10 min.

ness to trade accuracy for speed by indicating to some participants the expected task length.

Participants

Participants were 35 college and graduate students from Caltech recruited via the online system in the Social Science Experimental Laboratory. All participants received a \$5 show-up fee in addition to their earnings for performance as described below. The study was approved by the Caltech Committee for the Protection of Human Subjects.

Materials

The focal task was the random dot motion task, in each trial of which a hundred small white moving dots were displayed in random locations on a black background. A small number of these dots (“signal” dots) moved deterministically either all left or all right, while the rest moved in random directions. Twelve percent of the dots in all trials were signal dots (i.e., the coherence was always 12%). Participants had to choose, using the number keys “1” and “2” on the keyboard, which direction (left or right) the signal dots were moving in. They could respond at any time after the stimulus was first presented. The direction of coherent motion was determined with equal probability randomly across trials. Explicit instructions explaining the task including two comprehensive examples were provided before regular trials. The computerized experiment was programmed using

the Psychophysics Toolbox in MATLAB. The aperture (i.e., dot field) was square with side length 540 pixels. Dots had a 5-pixel diameter, a velocity of 1 pixel per frame (at roughly 60 fps), and a 20-frame lifetime. Intertrial intervals only comprised a fixation cross shown for 1.5 s, over which participants had no control.

Experimental Design

The full experimental design is depicted in Figure 1. Participants were divided into two conditions, and the key treatment was to instantiate the reference point at the outset of the dot motion task: the instructions in the treatment condition contained a line stating that the block “should take about 10 minutes to complete,” which constituted the experimental reference point. This was selected because it was a natural unit of time that was feasible to attain and slightly faster than the median completion time in pilot tests. More broadly, in the control condition, individuals completed two blocks of the random dot motion task.⁴ In the treatment condition, individuals completed a single block of the focal task (random dot motion) following one quick block of a different filler

⁴ The second block was intended to study the modulatory effects of prior experience on reference dependence. Because the results are suggestive but not clearly interpretable due to confounding effects of fatigue, I present the results in the Appendix for interested readers.

task (blurred image categorization).⁵ This served to stagger their start times on the focal task, avoiding confounds due to any external disruptions (such as loud noises) that could have spuriously mimicked a shared reference point. All blocks consisted of 100 trials and were separated by a 1-min break. Participants were paid \$0.05 for each correct answer and nothing for each wrong answer. Feedback was only provided as totals at the very end of the experiment. In what follows, I focus on the comparison between participants engaging in the dot motion task for the first time in both conditions. Although the treatment and control groups were not perfectly matched due to the auxiliary blocks, the data reveals no apparent differences between them before the reference point in terms of time expenditure or accuracy, as will be seen.⁶

During all tasks the real-time clock and trial number were displayed onscreen. At the end of the experiment, participants were asked how much they liked the task on a scale of 1 (*very little*) and 10 (*very much*). They were asked to remain in their seats until at least 30 min had passed from the start of the experiment before being paid but were allowed to browse the Internet in the meantime once they were finished. The marginal value of leisure was thus based on real leisure (Corgnet, Hernán-González, & Schniter, 2015). For laboratory timing reasons the experiment was set to end after 35 min and participants were informed of this. Eighteen people were assigned to the control condition and 17 people to the treatment condition. Two outlying participants from the control group were excluded from analysis since they took too much time and did not complete the full task, and so the two participants with the longest dot motion task times in the treatment group were excluded for balance (though the main results do not appreciably change if all participants are included).

Results

Reference dependence predicts that those who exceed the given reference point will finish working at a higher rate, at the expense of accuracy. Figure 2 shows the empirical distribution functions depicting the cumulative probabilities of stopping across the same groups, with 95% pointwise Wilson confidence bands. Those given the reference point stop at a higher rate, as indicated by a Cox proportional hazards

model ($HR = 2.45, p = .032$). In particular, the hazard rate is significantly elevated only after the reference point is hit ($HR_{\text{after}} = 3.38, p = .021$; $HR_{\text{before}} = 1.39, p = .620$). Further, the treatment reduced aggregate accuracy from 82.0% to 77.7% ($p = .003$, test for equality of proportions). Figure 3 shows the percent of correct answers in each group for individuals active before and after the reference point, with exact 95% confidence intervals. Before the reference point, both groups responded correctly about 80% of the time. Afterward, however, those given the reference point scored 7.7 percentage points lower than those who were not ($p = .015$). Although the control group appears to have increasing accuracy as the block proceeds which could be a reflection of training effects, this does not change the conclusion of the analysis. Variation in ability will moreover be accounted for via the drift rate parameter in the model estimation that follows.

Evidence of subjective displeasure comes from participants' postexperiment ratings of how much they liked the task on a scale from 1 (*very little*) to 10 (*very much*). Figure 4 and Figure 5 display the relationship between satisfaction ratings and task completion time, with 95% nonparametric bootstrap confidence intervals shown. Participants in the treatment group who spent longer than 10 min rated the task on average 3.3 points lower than those who finished quicker, a statistically significant difference according to a permutation test ($p = .044$). No such difference was found in the control group. Notwithstanding this result, it is possible that the drop in ratings simply has to do with

⁵ Participants spent between 4 min 16 s and 7 min 7 s on the filler block and had between 56% and 83% accuracy. They were faced with a sequence of images of animals (raccoons and porcupines), which were obscured using standard image processing filters. For each image they chose one of the two categories using the keys "1" and "2." The images were from a machine learning image set collected from an online image search. They were resized to approximately 200×300 pixels, converted to grayscale, and obscured using a 40-pixel range Gaussian blur.

⁶ This leaves open the possibility that other precipitating factors might contribute to the observed behavior. For example, the relative immediacy of leisure time in the treatment group could enhance the impact of the reference point. Investigating such factors and boundary conditions remains an interesting topic for further study, as with research that studies the role of salience in reference dependence (Bhatia & Golman, in press; Bordalo, Gennaioli, & Shleifer, 2012).

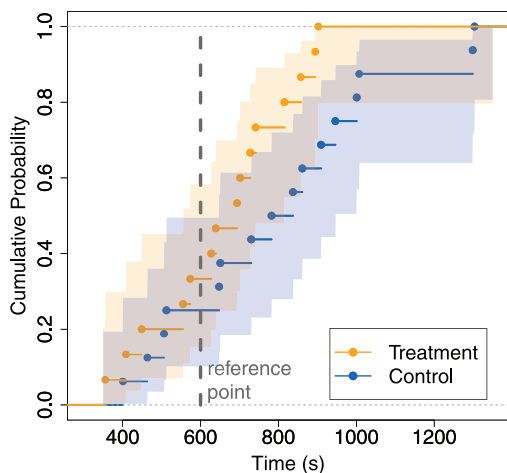


Figure 2. Empirical cumulative distribution function of completion time data. Shaded bands represent 95% pointwise Wilson confidence bands. See the online article for the color version of this figure.

participants putting in less effort and therefore giving noisier ratings, rather than necessarily signifying dissatisfaction. This is an alternative explanation which I cannot decisively rule out, and hence note as a caveat.

To more deeply characterize the data in terms of underlying preferences, I estimate the DDM parameters for each participant, and translate those into inferred subjective preference param-

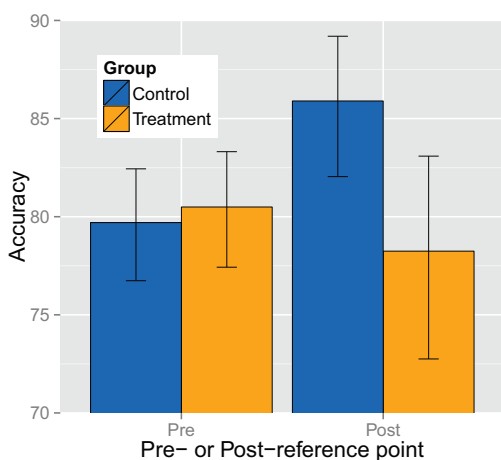


Figure 3. Accuracy rates across groups for participants active before and after reference point. Error bars represent exact 95% confidence intervals. See the online article for the color version of this figure.

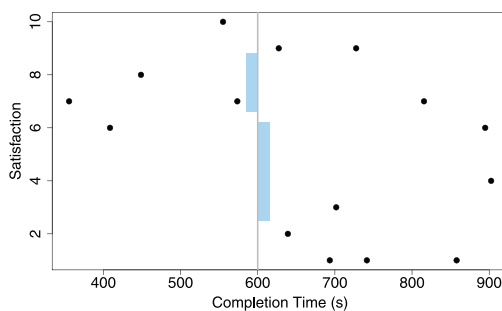


Figure 4. Subjective task satisfaction ratings on scale from 1 (low) to 10 (high) in treatment group. Blue (gray) bars represent 95% nonparametric bootstrap confidence intervals for people finishing before and after the reference point. See the online article for the color version of this figure.

eters. The drift rate A and decision threshold z are estimated using the EZ diffusion model (Wagenmakers, van Der Maas, & Grasman, 2007).⁷ This entails closed-form solutions for the parameters based only on the proportion of correct decisions (a) and the variance in response times for correct decisions ($var(t)$):

$$A = \text{sign}\left(a - \frac{1}{2}\right)c$$

$$\left\{ \frac{\text{logit}(a) \left[a^2 \text{logit}(a) - a \text{logit}(a) + a - \frac{1}{2} \right]}{\text{var}(t)} \right\}^{\frac{1}{4}},$$

$$z = \frac{2c^2 \text{logit}(a)}{A}.$$

where $\text{logit}(a) = \log\left(\frac{a}{1-a}\right)$. The properties of the DDM depend only on the ratios z/c and A/c rather than their absolute values so $c = .1$ is assumed in estimation as is standard practice.

I use the EZ diffusion approach to parameter estimation for two reasons. First, because of its simplicity, it often has superior power to detect experimental effects compared to other procedures, especially with a small number of trials (van Ravenzwaaij, Donkin, & Vandekerckhove, 2017). Second, its analytical tractability means

⁷ The nondescent time can also be estimated as the residual mean response time. I only briefly analyze this below.

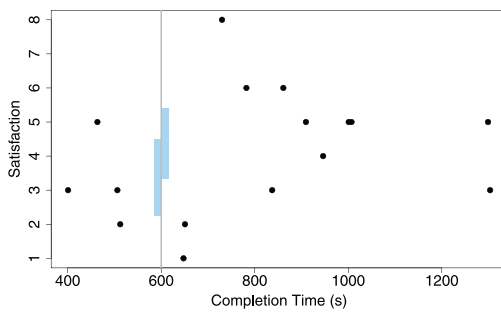


Figure 5. Subjective task satisfaction ratings on scale from 1 (*low*) to 10 (*high*) in control group. Blue (gray) bars represent 95% nonparametric bootstrap confidence intervals for people finishing before and after the reference point. See the online article for the color version of this figure.

that parameters are quick to numerically compute. Both of these properties will be needed particularly when estimating parameters using subsets of trials within a sliding temporal window. Moreover, the EZ model assumptions seem empirically reasonable in this dataset. Since simple flat threshold DDMs (and the EZ model) predict independence between time and accuracy, theoretical extensions such as collapsing thresholds or across-trial parameter variability are often motivated by negative correlations between these variables. In this dataset, logistic regressions (not reported) attempting to predict accuracy from time for each participant hold almost no predictive power; the time coefficient is not statistically significant for 85% of participants. Thus, basic models seem sufficient for the main goal of this article. This is perhaps to be expected on both theoretical and empirical grounds. Theoretically, a constant threshold is indeed optimal in the present task under the criterion used in this article (Fudenberg, Strack, & Strzalecki, *in press*), and empirically, meta-analyses typically find evidence primarily in favor of fixed thresholds (Hawkins, Forstmann, Wagenmakers, Ratcliff, & Brown, 2015). It is also not obvious how to incorporate reference dependence into complicated DDM variants which do not have clear optimizing foundations, making a simple model necessary for this application.⁸

The resulting estimates are shown in Figure 6 with 95% nonparametric bootstrap confidence intervals. The drift rate parameter is similar across groups both before ($p = .619$, across-

group permutation test) and after ($p = .709$) the reference point. Hence, changes in ability do not seem to explain the treatment effects.⁹ In contrast, the decision threshold drops after the reference point is hit for the group provided with the information ($p = .009$). This entails a shift toward speed and away from accuracy, all else equal. There were no statistically detectable differences in nonddecision time across groups either before ($p = .595$) or after ($p = .445$) the reference point. Thus, in line with many past studies varying speed-accuracy instructions, the only clear evidence of a treatment effect shows up in the decision threshold.

These parameter estimates can be translated into the opportunity cost of time using Expression (*). Analyzing the data on this level allows us to directly quantify how subjective preferences are affected by the treatment, accounting for individual variation in ability. I compute the value of time for participants in both groups based on their trials before and after the reference point was passed. Thus, the people represented in each group are the same across periods in this analysis. To keep noise low in these estimates, I include participants who spent enough time to face at least 20 trials after the reference point came into effect. The values resulting from the procedure on a \$/hour scale are shown in Figure 7 with 95% nonparametric bootstrap confidence intervals.

Table 1 contains the results of regressions predicting the values based on period (before vs. after reference point) and group (no-information control vs. reference point information treatment). As is apparent from the figure as well as the significant positive interaction between the two variables, the value of time rose dramatically only among those who were provided information and only after they exceeded the reference point. Before the reference point was passed, the value of time was the same regardless of whether groups were pro-

⁸ Nonetheless it must be noted that exact quantitative results are conditional on the model used and the assumptions it makes (such as constancy of parameters). These assumptions are not universally true and may require novel theoretical modifications depending on the application.

⁹ Drift rates do exhibit a mild and statistically significant increase likely due to improvement from experience as mentioned earlier ($p = .032$, paired t -test pooling both groups), but this does not alter the conclusion.

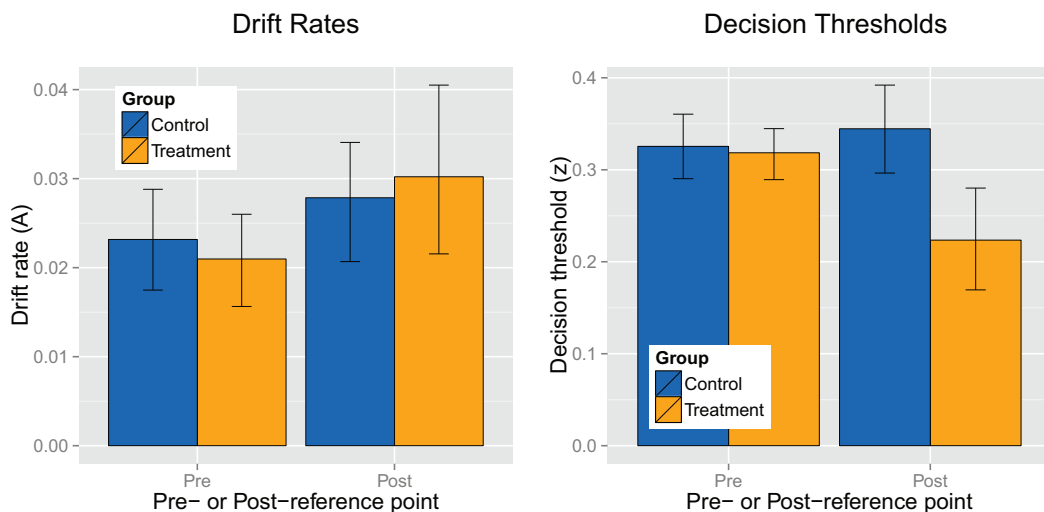


Figure 6. Mean estimated drift diffusion model parameters across groups for participants active before and after the reference point. Error bars represent 95% nonparametric bootstrap confidence intervals. See the online article for the color version of this figure.

vided information. After the reference point was passed, the control group’s preferences remained the same. Note that this is the case even though drift rates increased with experience; preference parameters naturally incorporate variation in ability by the way in which they are calculated, making them useful tools if one is interested in preferences specifically.

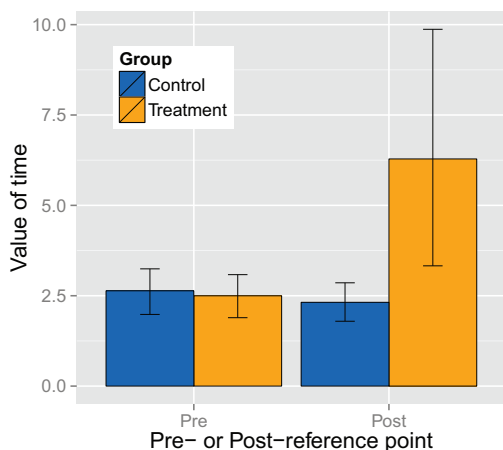


Figure 7. Mean estimated values of time across groups for participants active before and after the reference point. Error bars represent 95% nonparametric bootstrap confidence intervals. See the online article for the color version of this figure.

We can compare opportunity costs both between groups and within individuals to estimate the strength of loss aversion. Under the model specified earlier, the loss aversion parameter λ is given as the ratio between opportunity costs when loss aversion is and is not in effect. Note that because the value of a correct response factors in multiplicatively as seen in Expression (*), this is robust to certain assumptions about the utility of winning, including heterogeneous risk attitudes or psychological success bonuses. The within-person estimate is based on the ratio of opportunity costs after versus before the reference point for each individual in the treatment

Table 1
Effect of Reference Point on Value of Time

Variable	Dependent variable: Value of time	
	(1)	(2)
Constant	2.636 (.782)	3.187 (1.892)
Treatment	-.138 (1.219)	-2.149 (2.702)
After reference point	-.365 (1.106)	-.365 (1.141)
Treatment × After reference point	4.388 (1.723)	4.388 (1.778)
Individual fixed effects	No	Yes
Observations	34	34
R ²	.334	.646

Note. Standard errors are in parentheses.

group. The between-group estimate is based on the ratio of treatment to control group opportunity costs after the reference point. Although we cannot observe the behavior after the reference point of individuals who finish the task too quickly, if the loss aversion parameter is independent of the baseline opportunity cost, these assessments will provide unbiased estimates of loss aversion in the population. They agree with each other reasonably well. The mean within-person estimate is $\hat{\lambda}_w = 3.33$ with a 95% non-parametric bootstrap confidence interval of [1.34, 6.41], and the mean between-group estimate is $\hat{\lambda}_B = 2.87$ with a confidence interval of [1.45, 4.94]. These quantities summarize how intensely shortfalls of time are felt compared to surpluses. The figures here appear similar in magnitude to the most comparable figures from the few other time-based studies that exist. The best fitting models in Crawford and Meng's (2011) analysis yield values of 1.67, 2.31, and 2.89, while Abdellaoui and Kemel's (2014) find mean values of 2.54 and 3.80.¹⁰

The above analysis was conducted with a particular reference point in mind. However, the technique discussed can be used not only to measure the strength of loss aversion, but also to detect the location of its impact. Rather than dividing the data into regimes reflecting before and after the reference point, I estimate time-use preferences in a sliding window. Displayed in Figure 8 is the 10% trimmed mean of the value of time in each group based on a 3-min window for the period of time in which at least three individuals were active.¹¹ The inferred preferences are relatively stable and comparable in both groups until the reference point is hit. At that point, the mean value of time among the remaining participants given the reference point sharply rises. The horizontal bar above the data denotes periods with a statistically significant difference in group means at the 5% level according to a permutation test. Thus, reference points may be identified from the data itself. Intriguingly, there is only a little evidence of forward-looking sophistication, as the cost of time remains fairly stable in advance of reference point. However, the decision threshold does appear to be mildly lower among the treatment group starting from about halfway to the reference point (see the Appendix).

Discussion

In this article, I empirically investigated the effects of temporal reference dependence on deliberative time allocation. This was done within the context of a standard perceptual decision-making paradigm, the random dot motion task, in which behavior is known to be well-described by the DDM. By providing some participants with an expectation of how long a set of perceptual decision problems will take to complete, a benchmark was induced which, when exceeded, led people to shift their preferences toward speed over accuracy. Such people may also have experienced more subjective dissatisfaction as a result.

To directly quantify the treatment effects in terms of underlying preferences, I interpreted the decision threshold as a choice variable selected to trade off the decision maker's costs and benefits of time expenditure. This enabled me to infer the subjective value of time versus reward for each individual across trials. Within-individual and between-group analyses both revealed a roughly threefold increase in the cost of time after the reference point was exceeded. Moreover, by estimating preferences based on data within a sliding window, the reference point could be recovered without assuming its existence.

The preference-based analysis used in this article rests on two theoretical assumptions: that the costs and benefits of time expenditure are exclusively captured by the decision threshold, and that the decision threshold reflects the optimal balance between costs and benefits. The first assumption seems borne out by the data, as the only parameter modulated by the treatment appeared to be the decision threshold, while drift rate and nondecision time did not show

¹⁰ It must be noted that Crawford and Meng's (2011) parameter includes an extra coefficient reflecting other aspects of reference dependence, and Abdellaoui and Kemel's (2014) study involves framed gambles over amounts of time that participants were made to spend in a room later without entertainment, and are therefore not perfectly comparable to my estimates. De Borger and Fosgerau (2008) quantitatively assessed loss aversion from hypothetical travel time choices but do not estimate a comparable parameter.

¹¹ I exclude each individual's 100th trial and maximum estimated value of time, due to a sizable drop in performance specific to the last trial, and to ensure the results are not driven by other outliers.

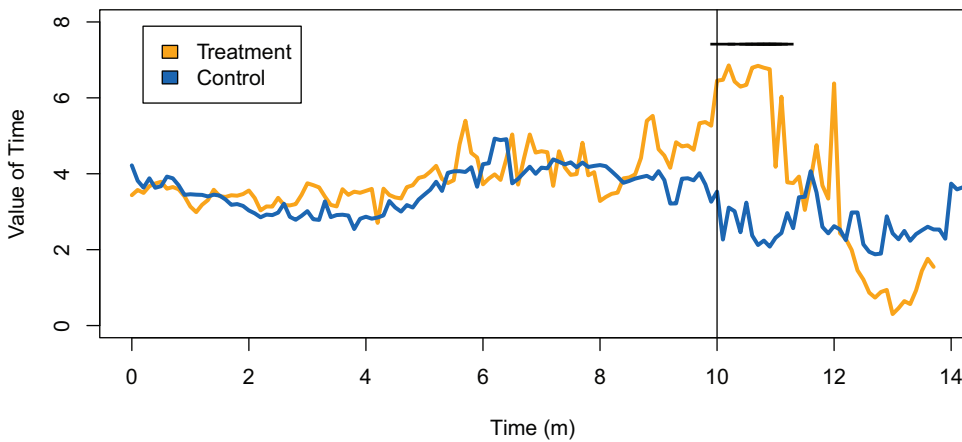


Figure 8. Moving average of estimated values of time across groups. Horizontal bar represents periods in which values are statistically significantly different across groups at the 5% level according to a permutation test. See the online article for the color version of this figure.

discernible differences across groups. This assumption is moreover consistent with many other studies in this context. For example, participants in Experiment 2 of Palmer et al. (2005) were instructed to aim for various mean response times (0.5 s, 1 s, or 2 s), and this manipulation only yielded lower thresholds with no clear changes in other parameters (though see Zhang & Rowe, 2014). Furthermore, neural evidence indicates that strategic control of the speed–accuracy trade-off appears regulated by downstream decision-related regions (such as presupplementary motor area, striatum, and dorsolateral prefrontal cortex), as opposed to early sensory areas which might be expected to reflect parameters other than the decision threshold (Bogacz, Wagenmakers, Forstmann, & Nieuwenhuis, 2010; Forstmann et al., 2008, 2010; Ivanoff et al., 2008; van Veen et al., 2008; Wenzlaff, Bauer, Maess, & Heekeren, 2011).

Although the second assumption cannot be directly tested here, to a limited extent the proof is in the pudding. This study and others have found that the estimated decision threshold and its neural correlates qualitatively respond to incentives and instructions as predicted by principles of optimality (Domenech & Dreher, 2010; Gluth, Rieskamp, & Büchel, 2013; Green, Biele, & Heekeren, 2012; Hanks, Kiani, & Shadlen, 2014). The most relevant quantitative evidence comes from Drugowitsch, Moreno-Bote, Churchland,

Shadlen, and Pouget (2012) who used similar methods to estimate within-trial cost functions for monkeys and humans. They found that inferred costs seemed to exhibit certain rational properties. For example, in the experimental protocol faced by the monkeys, rewards for quick responses were delayed to a minimum reward time of approximately 1 s (whereas any later decisions resulted in immediate reward). Hence, there was no cost to continued accumulation early on, and indeed, inferred costs were close to zero before the minimum reward time. Thus, internal costs appear sensitive to the nature of decision problems at least on short timescales.

Countless studies have influenced deliberation by manipulating objective features of the decision task such as monetary incentives or trial timing (Heitz, 2014). Fewer have altered subjective perceptions of the task, and almost none have quantified the effects of their manipulations in terms of preferences. This quantification is important for both predicting behavior precisely and understanding the internal motivations that contribute to deliberative decision making. By more firmly linking the DDM to its cost-benefit foundations, we might hopefully gain further insight into perceptual and economic decisions as well as generate deeper hypotheses regarding the neural implementation of individual deliberation. In separate paradigms, the basal ganglia have been associated

with subjective value representations including loss aversion (Canessa et al., 2013; Tom, Fox, Trepel, & Poldrack, 2007) and modulation of the decision threshold (Forstmann et al., 2008, 2010; Ivanoff et al., 2008; van Veen et al., 2008). Thus, there are many possible avenues for exploring the interaction of deliberation with reward processing (e.g., Frank et al., 2015).

Future research might consider other insights on subjective preferences in time use. For instance, reference dependence in the monetary domain could be studied by setting expectations about monetary earnings, as has been done in many different economic paradigms (Abeler, Falk, Goette, & Huffman, 2011; Ericson & Fuster, 2011; Gill & Prowse, 2012; Heffetz & List, 2014; Smith, 2012). Like the present study, this may induce participants to work less hard once they have reached the target. More broadly, the value of time is relatively ambiguous and lacks stable benchmarks—a dollar is a dollar, but a moment of life is less interchangeable with any other moment. Research in psychology and marketing accordingly reveals a great deal of flexibility in the appraisal and allocation of time (Okada & Hoch, 2004) and shows that people often allocate time more heuristically than money (Saini & Monga, 2008). Such perspectives may prove important to understanding how deliberation unfolds over time.

References

- Abdellaoui, M., & Kelkel, E. (2014). Eliciting prospect theory when consequences are measured in time units: "Time is not money." *Management Science*, *60*, 1844–1859.
- Abeler, J., Falk, A., Goette, L., & Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, *101*, 470–492.
- Arrow, K. J., Blackwell, D., & Girshick, M. A. (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica*, *17*, 213–244.
- Band, G. P. H., Ridderinkhof, K. R., & van der Molen, M. W. (2003). Speed-accuracy modulation in case of conflict: The roles of activation and inhibition. *Psychological Research*, *67*, 266–279.
- Bartling, B., Brandes, L., & Schunk, D. (2015). Expectations as reference points: Field evidence from professional soccer. *Management Science*, *61*, 2646–2661.
- Bell, D. E. (1985). Disappointment in decision making under uncertainty. *Operations Research*, *33*, 1–27.
- Bhatia, S., & Golman, R. (in press). Attention and reference dependence. *Decision*.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*, 700–765.
- Bogacz, R., Wagenmakers, E.-J., Forstmann, B. U., & Nieuwenhuis, S. (2010). The neural basis of the speed-accuracy tradeoff. *Trends in Neurosciences*, *33*, 10–16.
- Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience in experimental tests of the endowment effect. *American Economic Review*, *102*, 47–52.
- Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review*, *100*, 432–459.
- Canessa, N., Crespi, C., Motterlini, M., Baud-Bovy, G., Chierchia, G., Pantaleo, G., . . . Cappa, S. F. (2013). The functional and structural neural basis of individual differences in loss aversion. *Journal of Neuroscience*, *33*, 14307–14317.
- Corgnet, B., Hernán-González, R., & Schniter, E. (2015). Why real leisure really matters: Incentive effects on real effort in the laboratory. *Experimental Economics*, *18*, 284–301.
- Crawford, V. P., & Meng, J. (2011). New York City cab drivers' labor supply revisited: Reference-dependent preferences with rational-expectations targets for hours and income. *American Economic Review*, *101*, 1912–1932.
- De Borger, B., & Fosgerau, M. (2008). The trade-off between money and travel time: A test of the theory of reference-dependent preferences. *Journal of Urban Economics*, *64*, 101–115.
- de Hollander, G., Labruna, L., Sellaro, R., Trutti, A., Colzato, L. S., Ratcliff, R., . . . Forstmann, B. U. (2016). Transcranial direct current stimulation does not influence the speed-accuracy tradeoff in perceptual decision-making: Evidence from three independent studies. *Journal of Cognitive Neuroscience*, *28*, 1283–1294.
- Domenech, P., & Dreher, J.-C. (2010). Decision threshold modulation in the human brain. *Journal of Neuroscience*, *30*, 14305–14317.
- Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica*, *30*, 412–431. (Originally published in 1869)
- Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, *32*, 3612–3628.

- Edwards, W. (1965). Optimal strategies for seeking information: Models for statistics, choice reaction times, and human information processing. *Journal of Mathematical Psychology*, 2, 312–329.
- Eliasz, K., & Spiegler, R. (2013). Reference-dependence and labor-market fluctuations. *NBER Macroeconomics Annual*, 28, 159–200.
- Ericson, K., Marzilli, M., & Fuster, A. (2011). Expectations as endowments: Evidence on reference-dependent preferences from exchange and valuation experiments. *Quarterly Journal of Economics*, 126, 1879–1907.
- Forstmann, B. U., Anwander, A., Schäfer, A., Neumann, J., Brown, S., Wagenmakers, E.-J., . . . Turner, R. (2010). Cortico-striatal connections predict control over speed and accuracy in perceptual decision making. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 15916–15920.
- Forstmann, B. U., Dutilh, G., Brown, S., Neumann, J., Von Cramon, D. Y., Ridderinkhof, K. R., & Wagenmakers, E.-J. (2008). Striatum and pre-SMA facilitate decision-making under time pressure. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 17538–17542.
- Forstmann, B. U., Ratcliff, R., & Wagenmakers, E.-J. (2016). Sequential sampling models in cognitive neuroscience: Advantages, applications, and extensions. *Annual Review of Psychology*, 67, 641–666.
- Frank, M. J., Gagne, C., Nyhus, E., Masters, S., Wiecki, T. V., Cavanagh, J. F., & Badre, D. (2015). fMRI and EEG predictors of dynamic decision parameters during human reinforcement learning. *Journal of Neuroscience*, 35, 485–494.
- Fudenberg, D., Strack, P., & Strzalecki, T. (in press). Speed, accuracy, and the optimal timing of choices. *American Economic Review*.
- Gill, D., & Prowse, V. (2012). A structural analysis of disappointment aversion in a real effort competition. *American Economic Review*, 102, 469–503.
- Gluth, S., Rieskamp, J., & Büchel, C. (2013). Classic EEG motor potentials track the emergence of value-based decisions. *Neuroimage*, 79, 394–403.
- Gold, J. I., & Shadlen, M. N. (2002). Banburismus and the brain: Decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36, 299–308.
- Gold, J. I., & Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535–574.
- Green, N., Biele, G. P., & Heekeren, H. R. (2012). Changes in neural connectivity underlie decision threshold modulation for reward maximization. *Journal of Neuroscience*, 32, 14942–14950.
- Gul, F. (1991). A theory of disappointment aversion. *Econometrica*, 59, 667–686.
- Hale, D. J. (1969). Speed-error tradeoff in a three-choice serial reaction task. *Journal of Experimental Psychology*, 81, 428–435.
- Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science*, 274, 427–430.
- Hanks, T., Kiani, R., & Shadlen, M. N. (2014). A neural mechanism of speed-accuracy tradeoff in macaque area LIP. *eLife*, 3, e02260.
- Hawkins, G. E., Forstmann, B. U., Wagenmakers, E.-J., Ratcliff, R., & Brown, S. D. (2015). Revisiting the evidence for collapsing boundaries and urgency signals in perceptual decision-making. *Journal of Neuroscience*, 35, 2476–2484.
- Heffetz, O., & List, J. A. (2014). Is the endowment effect an expectations effect? *Journal of the European Economic Association*, 12, 1396–1422.
- Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, 8, 150.
- Herz, D. M., Tan, H., Brittain, J.-S., Fischer, P., Cheeran, B., Green, A. L., . . . Brown, P. (2017). Distinct mechanisms mediate speed-accuracy adjustments in cortico-subthalamic networks. *eLife*, 6, e21481.
- Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 4, 11–26.
- Howell, W. C., & Kreidler, D. L. (1963). Information processing under contradictory instructional sets. *Journal of Experimental Psychology*, 65, 39–46.
- Ivanoff, J., Branning, P., & Marois, R. (2008). fMRI evidence for a dual process account of the speed-accuracy tradeoff in decision-making. *PLoS ONE*, 3(7), e2635.
- Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Amsterdam, the Netherlands: Elsevier.
- Kőszegi, B., & Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121, 1133–1165.
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13, 1292–1298.
- Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology*, 3, 193.
- Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 13852–13857.
- LaBerge, D. (1962). A recruitment theory of simple behavior. *Psychometrika*, 27, 375–396.

- Laming, D. R. J. (1968). *Information theory of choice-reaction times*. London, UK: Academic Press.
- Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies*, *53*, 271–282.
- Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making*, *5*, 437–449.
- Okada, E. M., & Hoch, S. J. (2004). Spending time versus spending money. *Journal of Consumer Research*, *31*, 313–323.
- Osman, A., Lou, L., Muller-Gethmann, H., Rinke-nauer, G., Mattes, S., & Ulrich, R. (2000). Mechanisms of speed–accuracy tradeoff: Evidence from covert motor processes. *Biological Psychology*, *51*, 173–199.
- Pagel, M. (2017). Expectations-based reference-dependent life-cycle consumption. *Review of Economic Studies*, *84*, 885–934.
- Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, *5*, 1.
- Philiastides, M. G., Ratcliff, R., & Sajda, P. (2006). Neural representation of task difficulty and decision making during perceptual categorization: A timing diagram. *Journal of Neuroscience*, *26*, 8965–8975.
- Pike, A. R. (1966). Stochastic models of choice behaviour: Response probabilities and latencies of finite Markov chain systems. *British Journal of Mathematical and Statistical Psychology*, *19*, 15–32.
- Pope, D. G., & Schweitzer, M. E. (2011). Is Tiger Woods loss averse? Persistent bias in the face of experience, competition, and high stakes. *American Economic Review*, *101*, 129–157.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Ratcliff, R., Cherian, A., & Segraves, M. (2003). A comparison of macaque behavior and superior colliculus neuronal activity to predictions from models of two-choice decisions. *Journal of Neurophysiology*, *90*, 1392–1407.
- Ratcliff, R., Philiastides, M. G., & Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 6539–6544.
- Ratcliff, R., Smith, P. L., Brown, S. D., & McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, *20*, 260–281.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347–356.
- Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological Review*, *111*, 333–367.
- Rinkenauer, G., Osman, A., Ulrich, R., Müller-Gethmann, H., & Mattes, S. (2004). On the locus of speed-accuracy trade-off in reaction time: Inferences from the lateralized readiness potential. *Journal of Experimental Psychology: General*, *133*, 261–282.
- Saini, R., & Monga, A. (2008). How I decide depends on what I spend: Use of heuristics is greater for time than for money. *Journal of Consumer Research*, *34*, 914–922.
- Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: Seeing and deciding. *Proceedings of the National Academy of Sciences of the United States of America*, *93*, 628–633.
- Shadlen, M. N., & Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area LIP) of the rhesus monkey. *Journal of Neurophysiology*, *86*, 1916–1936.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701.
- Smith, A. (2012). *Lagged beliefs and reference-dependent preferences*. New York, NY: Mimeo.
- Smith, P. L., & Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends in Neurosciences*, *27*, 161–168.
- Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*, 652–654.
- Stone, M. (1960). Models for choice-reaction time. *Psychometrika*, *25*, 251–260.
- Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science*, *315*, 515–518.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.
- van der Lubbe, R. H. J., Jaśkowski, P., Wauschkuhn, B., & Verleger, R. (2001). Influence of time pressure in a simple response task, a choice-by-location task, and the Simon task. *Journal of Psychophysiology*, *15*, 241–255.
- van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The EZ diffusion model provides a powerful test of simple empirical effects. *Psychonomic Bulletin & Review*, *24*, 547–556.

- van Veen, V., Krug, M. K., & Carter, C. S. (2008). The neural and computational basis of controlled speed-accuracy tradeoff during task performance. *Journal of Cognitive Neuroscience*, *20*, 1952–1965.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, *13*, 37–58.
- Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition*, *32*, 1206–1220.
- Wagenmakers, E.-J., van Der Maas, H. L. J., & Grasman, R. P. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3–22.
- Wald, A. (1947). *Sequential analysis*. New York, NY: John Wiley & Sons.
- Wald, A., & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Annals of Mathematical Statistics*, *19*, 326–339.
- Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, *36*, 955–968.
- Wenzlaff, H., Bauer, M., Maess, B., & Heekeren, H. R. (2011). Neural characterization of the speed-accuracy tradeoff in a perceptual decision-making task. *Journal of Neuroscience*, *31*, 1254–1266.
- Zhang, J., & Rowe, J. B. (2014). Dissociable mechanisms of speed-accuracy tradeoff during visual perceptual learning are revealed by a hierarchical drift-diffusion model. *Frontiers in Neuroscience*, *8*, 69.

Appendix

Analysis of Intertemporal Sophistication

Decision makers with temporal deadlines such as reference points could in principle alter their behavior in a forward-looking manner, speeding up in advance of the deadline to improve their chances of timely completion. The earlier analysis reveals little evidence of this, although Figure A1 which plots the continuously-estimated parameters¹² hints at a suggestive lowering of the decision threshold around the 5-min mark. This is intriguing as 5 min is half-way to the reference point and could indicate a temporal benchmark that is used for pacing.

Forward-looking pacing may be easier when one has experience with the task. Figure A2 shows the distribution of completion times for each dot motion block, including for participants who were provided the reference point

their second time around. It indicates their completion times cluster right before the reference point. Figure A3 shows the moving average of the cost of time for this group, which is relatively elevated. This could be a result of either sophistication or fatigue. Interestingly, the cost of time peaks around the 5-minute mark. Whether this reflects a temporal benchmark or is simply a coincidence is unknown.

¹² It should be noted that the mean parameter values get noisier towards the end of the time period because the number of participants remaining (i.e., those who have not finished the task at each point in time) declines. The oscillation in nondecision time appears to be the result of erratic estimates for a single subject.

(Appendix continues)

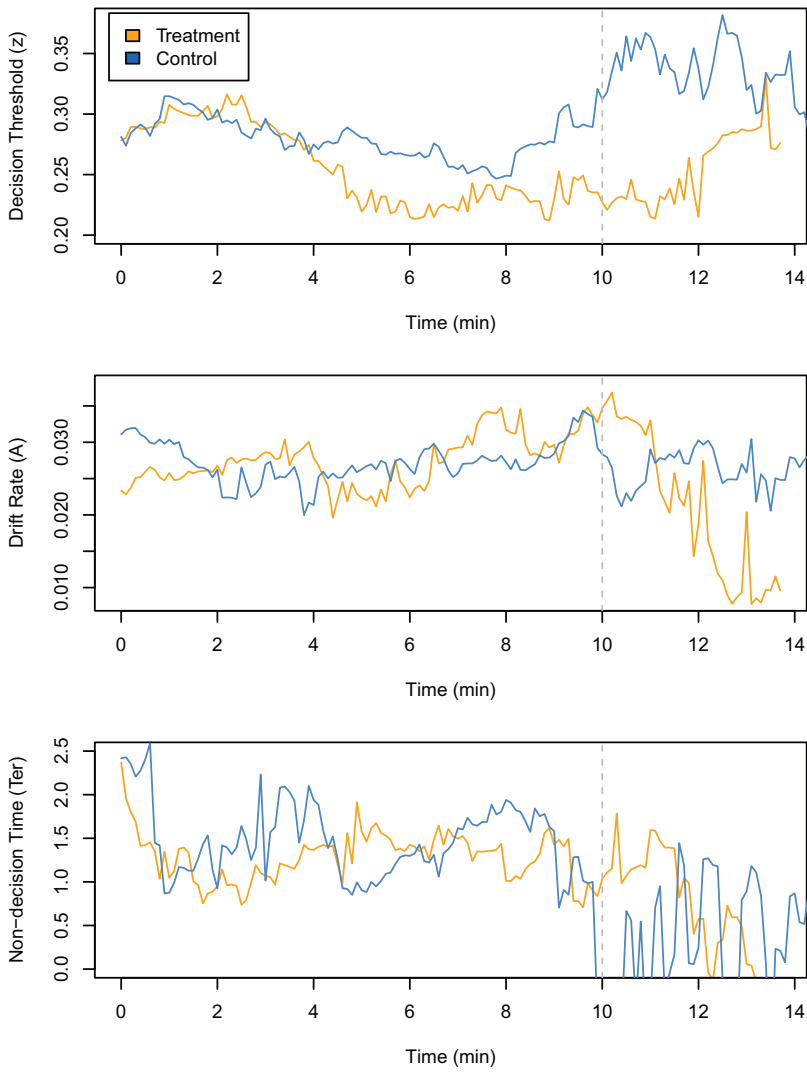


Figure A1. Moving average of estimated drift diffusion model parameters across groups. Blue (dark gray) and orange (light gray) represent control and treatment groups. See the online article for the color version of this figure.

(Appendix continues)

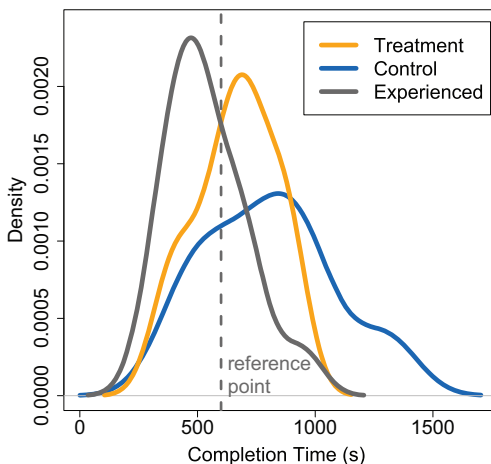


Figure A2. Kernel density estimate of completion time data with experienced participants. See the online article for the color version of this figure.

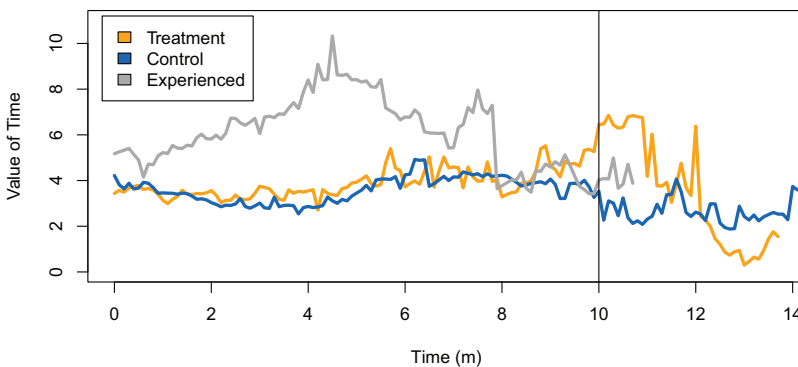


Figure A3. Moving average of estimated values of time with experienced participants. See the online article for the color version of this figure.

Received May 15, 2018
 Revision received July 4, 2018
 Accepted July 12, 2018 ■